Available online at :
https://ejournal.unsrat.ac.id/v3/index.php/IJIDS/index

IJIDS
(Indonesian Journal of Intelligence Data Science)

Sentimen Analisis terhadap implementasi program Kampus Merdeka dengan Algoritma *Support Vector Machine* (SVM) dan Naïve Bayes

Handira Leo Putra Sinaga*[1] , Winsy C. D Weku[2] , Charles E. Mongi[3]
[1,2]Program Studi Sistem Informasi, FMIPA, UNSRAT
e-mail: dsinaga9@gmail.com, winsy_weku@unsrat.ac.id, charlesmongi@ymail.com

ARTICLE INFO

History of the article:
Received March 1, 2022
Revised March 21, 2022
Accepted January 12, 2022

Keywords:
3 to 5
Keywords:
Analisis sentimen, svm, naive bayes, mbkm

Correspondece:
E-mail:
dsinaga9@gmail.com

ABSTRAKSI

Program Kampus Merdeka merupakan program pemerintah untuk meningkatkan kualitas pendidikan tinggi di Indonesia. Untuk implementasinya, program ini menimbulkan berbagai macam tanggapan dan perasaan dari masyarakat, termasuk dari mahasiswa dan alumni perguruan tinggi. Oleh karena itu, penting untuk menganalisis perasaan dan opini tersebut untuk memahami bagaimana implementasi program ini diterima oleh masyarakat menggunakan analisis sentimen. Data akan diambil dari cuitan Twitter dengan menggunakan #kampusmerdeka. Pada penelitian ini, analisis sentimen akan dilakukan terhadap data teks cuitan Twitter yang berisi tanggapan terkait implementasi program Kampus Merdeka. Algoritma *Support Vector Machine* (SVM) dan Naive Bayes akan digunakan sebagai metode untuk melakukan analisis sentimen tersebut. Dengan menggunakan pembagian data 80:20 untuk data pelatihan dan data uji. Didapatkan akurasi untuk SVM sebesar 98%, dan Naive Bayes sebesar 85%. Keduanya menunjukkan kemampuan yang baik dalam memprediksi sentimen secara keseluruhan pada dataset yang digunakan.

Kata Kunci: Analisis sentimen, svm, naive bayes, mbkm

PENDAHULUAN

Program Kampus Merdeka merupakan program pemerintah untuk meningkatkan kualitas pendidikan tinggi di Indonesia. Untuk implementasinya, program ini menimbulkan berbagai macam tanggapan dan perasaan dari masyarakat, termasuk dari mahasiswa dan alumni perguruan tinggi. Oleh karena itu, penting untuk menganalisis perasaan dan opini tersebut untuk memahami bagaimana implementasi program ini diterima oleh masyarakat [1], [2].

Analisis sentimen dapat digunakan untuk menentukan perasaan positif, negatif, atau netral yang terkandung dalam teks. Pada penelitian ini, analisis sentimen akan dilakukan terhadap data teks yang berisi tanggapan terkait implementasi program Kampus Merdeka. Algoritma Support Vector Machine (SVM) dan Naive Bayes akan digunakan sebagai metode untuk melakukan analisis sentimen tersebut [3]. Kedua algoritma ini memiliki kelebihan dan kekurangan masingmasing, sehingga perbandingan antara kedua algoritma tersebut akan dilakukan untuk menentukan metode mana yang lebih baik digunakan dalam konteks ini [4], [5].

Penelitian terdahulu yang telah dilakukan terkait Sentimen Analis adalah sebagai berikut, Analisis Sentimen Kebijakan Magang Merdeka Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM,

Pengumpulan dan persiapan dataset dimulai dengan seleksi fitur, menghilangkan duplikasi dan seleksi tweet, kemudian dilakukan pre-processing . Hasil penelitian ini menghasilkan model LSTM yang telah dilatih dari dataset 658 tweet dengan nilai akurasi terbaik di 80,42%. Analisis sentimen program Magang Merdeka dari tweet pengguna didominasi oleh perasaan "bingung" yaitu 39,51%, kemudian disusul oleh perasaan "senang" yaitu 16,26%, perasaan "sedih" yaitu 15,80%, perasaan "marah" yaitu 13,98%, perasaan "takut" yaitu 7,29%, dan perasaan "terkejut" yaitu 7,14% [6]. Lalu ada Analisis Sentimen dengan Naive Bayes terhadap Komentar Aplikasi Tokopedia dimana dalam penelitian ini berupa kategori positif dan negatif. Penelitian ini menggunakan metode Naive Bayes untuk menghasilkan sentimen positif dan negatif terhadap komentar pengguna aplikasi Tokopedia di Playstore. Pengujian berdasarkan nilai class negative, class positive, recall, dan accuracy pada analisis sentimen. dengan nilai accuracy performance yang baik sebesar 97,13%, dengan nilai precision 1 Sementara pada Class Recall dihasilkan nilai 95,49% (positive class: negative). Dan nilai AUC 0,980 [7]. Lalu ada juga penelitian tentang Perbandingan Metode Algoritma Naïve Bayes Classifier Dan Support Vector Machine Pada Kasus Analisis Sentimen Terhadap Data Vaksin Covid-19 Di Twitter, Penelitian ini mengkaji kinerja Naïve Bayes Classifier (NBC) dan Support Vector Machine (SVM) dengan menambahkan teknik TF-IDF (Term FrequencyInverse Document Frequency) yang bertujuan untuk memberikan bobot pada hubungan kata (term) sebuah dokumen. Kemudian melakukan splitting data yaitu membagi data training 80% dan data testing 20%. Hasil klasifikasi data tweet vaksin Covid-19 menggunakan algoritma Naïve Bayes Classifier mendapatkan nilai accuracy sebesar 81%, precision sebesar 80%, recall sebesar 99%, dan f1-score sebesar 89%, Sedangkan untuk algoritma Support Vector Machine mendapatkan nilai accuracy sebesar 87%, precision sebesar 88%, recall sebesar 96%, dan f1- score sebesar 92% [8].

Penelitian ini dilakukan dengan menggunakan data teks yang diambil dari cuitan komentar dari Twitter. Alasan penggunaan twitter oleh peneliti sebagai media untuk melihat sentimen atau opini publik disebabkan angka penggunaan twitter di Indonesia sangat tinggi dengan menduduki peringkat ketiga sebagai pengguna media Twitter terbanyak.

Secara keseluruhan, penelitian ini memiliki tujuan untuk mengetahui perasaan masyarakat terhadap implementasi program Kampus Merdeka dan membandingkan kemampuan dua algoritma SVM dan Naive Bayes dalam melakukan analisis sentimen. Penelitian ini diharapkan dapat memberikan wawasan bagi pemerintah dan pihak terkait tentang bagaimana implementasi program Kampus Merdeka diterima oleh masyarakat dan apa saja yang menjadi perhatian utama dalam implementasi program tersebut.

METODE PENELITIAN

1. Tahapan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining dan pembelajaran mesin (machine learning) [9], [10] untuk menganalisis sentimen dari tanggapan mahasiswa, dosen, atau stakeholders terkait implementasi program Kampus Merdeka.

2. Sumber Data:

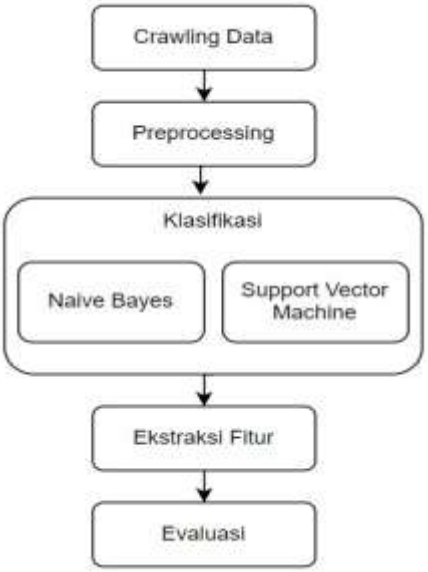
Sumber data yang digunakan dalam riset penelitian ini berupa data sekunder. Data sekunder merupakan suatau data yang diambil dari beberapa literatur dan sumber dari media lain [11], [12]. Data teks tidak terstruktur yang diperoleh dari berbagai literatur berikut ini:

- Media sosial (Twitter, Instagram, Facebook) menggunakan API.
- Survei online terhadap peserta program Kampus Merdeka.

- Forum diskusi (Kaskus, Reddit, atau platform edukasi).
- Ulasan di platform kampus (SIAKAD, LMS).

3. Tahapan Penelitian

Tahapan penelitian yang dilakukan pada analisis sentimen adalah *crawling* data, preprocessing, ekstraksi fitur, klasifikasi,dan Evaluasi [13]. Data di *crawling* dari data twitter dengan cara mengambil API komentar Twitter.



Gambar 1. Tahapan Penelitian

4. Pengambilan Data

Data di crawling dari data twitter dengan cara mengambil API komentar Twitter dengan merequset token API dari *Twitter Developer*.dan data diambil dari tanggal 1 Januari 2022. Data *dicrawling* menggunakan *hashtag* #mbkm.

Gambar 2. Tahapan Proses pengambilan Data Twitter

Hasil dari *crawling* data disimpan dalam file dan kemudian dilakukan labelling untuk menentukan pendapat atau pandangan dari komentar yang diambil. Pada proses labelling dibedakan menjadi 3 kelas, yaitu kelas positif, kelas negatif, dan kelas netral. contoh dari data yang didapat dan hasil dari labeling seperti terlihat pada Tabel 1.

Tabel 1. Label pada komentar

Komentar	**Label**
Program MBKM membawa Mahasiswa Go Internasional	Positif
KKN MBKM UNS Belajar Jadi Wirausahawan Ulet dari Perajin Ban Bekas di Miri	Positif
Kampus lain cuma upload laporan seminggu sekali, di UM kalo MBKM harus buat laporan tiap hari terus diakhir disusun kaya skripsi. Capek bgt mata I berat	Negatif
paling males kalau udah tugas kelompok mbkm. ini aku saran malah dicuekin. fuck, kalau merasa bisa ngurusin sendiri ya silakan. capek.	Negatif
Sampe tau kabar dari temen ada program MBKM iisma tapi sempet kecewa karena awalnya cuma dibuka buat S1 bukan vokasi ? untung awal tahun 2022 pas msib dapat kabar dari temen kalo iisma bukan jalur vokasi. Tanpa pikir 2x langsung minta restu dan alhamdulillah mama full support	Netral

3. Preprocessing

Langkah pembersihan data sebelum diklasifikasikan. Preprocessing menghindari pengocokan data, data tidak sempurna dan data yang tidak konsisten. Pengolahan data yaitu pembersihan data mentah yang diperoleh selama proses crawling agar siap untuk digunakan pada langkah selanjutnya seperti cleaning, remove stopword, tokenization, stemming [14], [15].

4. Ekstraksi Fitur

Proses ekstraksi fitur pada komentar yaitu dengan *Term Frequency* dan TFIDF (*Term Frequency-Inverse Document Frequency*) [16]. Langkah-langkah dalam tahap ini adalah sebagai berikut:

a. Hitung *Term Frequency, Term Frequency* yaitu menghitung kemunculan suatu term pada suatu corpus.
b. Hitung TF-IDF (*Term Frequency-Inverse Document Frequency*) Langkah pertama perhitungan TF-IDF adalah menghitung *Inverse document frecuency* (idf). Setelah idf diketahui, maka langkah berikutnya adalah mencari nilai bobot TF-IDF, dengan cara mengalikan nilai *term frequency* dengan nilai *inverse document frequency* untuk masing-masing term.

5. Klasifikasi

Proses untuk mengkategorikan komentar yang sentimennya tidak diketahui. Proses klasifikasi menggunakan algoritma SVM dan Naive Bayes yang terbagi menjadi dua set data yaitu training dan data testing. Kemudian data akan diklasifikasikan berdasarkan ketetapan Confusion Matrix.

6. Evaluasi

Evaluasi tes dilakukan untuk mencari keakuratan hasil model terbaik. Evaluasi menggunakan *confusion matrix*,dimana *Confusion Matrix* merupakan metode evaluasi yang dapat digunakan untuk menghitung kinerja atau tingkat kebenaran dari proses klasifikasi. Ada empat istilah yang merupakan representasi hasil proses klasifikasi pada *Confusion Matrix* yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

HASIL DAN PEMBAHASAN

1. Ekstraksi Fitur

Pada proses ekstraksi fitur, tahap pertama yang dilakukan setelah tokenisasi yaitu mengubah dataset ke dalam sebuah representasi vector dengan menggunakan library python yang bernama CountVectorizer. Setiap baris pada Vector mewakili baris setiap data komentar yang menjadi dataset. Kemudian setiap kolomnya merupakan seluruh kata. Sebagai contoh terdapat 3 komentar yang berhasil didapatkan, di antaranya :

(D1) “berharap banget keterima msib magang batch amin”

(D2) “jadi lebih ke msib versi international gak sih keren banget”

(D3) “pengen ikut msib magang tapi ada pkl”

Setelah dilakukan tahapan-tahapan sebelumnya terdapat beberapa kata baku yaitu “keren”, “banget”, “ikut”, dan “magang”. Pada tahap selanjutnya tiap dokumen diwujudkan sebagai sebuah vektor dengan elemen, jika terdapat kata yang bersangkutan di dalam dokumen maka diberi nilai 1, jika tidak ada diberi nilai 0.

Sebagai contoh dapat dilihat pada Tabel 2.

Tabel 2. Pembuatan *Word Vector*

	Keren	**Banget**	**Ikut**	**Magang**
(D1)	0	1	0	0
(D2)	1	1	0	0
(D3)	0	0	1	1

Data yang sudah menjadi word vector kemudian dihitung menggunakan rumus TF-IDF sehingga menghasilkan word vector dengan nilai yang sudah terbobot. Adapun TF (Term Frequency) adalah frekuensi dari kemunculan sebuah term dalam dokumen yang bersangkutan, sedangkan IDF (Inverse Document Frequency) merupakan sebuah perhitungan dari bagaimana term didistribusikan secara luas pada koleksi dokumen yang bersangkutan. Proses pembobotan kata dilakukan dengan menghitung TF (Term Frequency) terlebih dahulu. Adapaun contohnya dapat dilihat pada Tabel 3.

Tabel 3. TF *(Term Frequency)*

	(D1)	**(D2)**	**(D3)**
Keren	0	1	0
Banget	1	1	0
Ikut	0	0	1
Magang	0	0	1

Setelah terbentuk TF (*Term Frequency*) selanjutnya menentukan DF *(Document Frequency)* yaitu banyaknya term (T) muncul dalam semua dokumen.

Adapun hasilnya dapat dilihat pada Tabel 4.

Tabel 4. Tabel DF *(Document Frequency)*

T (Term)	***DF(Document Frequency)***
Keren	1
Banget	2
Ikut	1
Magang	1

Kemudian menghitung nilai IDF (*Inverse Document Frequency*) dengan menghitung nilai log dari hasil D (jumlah dokumen) dibagi dengan nilai DF (*Document Frequency*). Adapun hasilnya dapat dilihat pada Tabel 5.

Tabel 5. IDF *(Inverse Document Frequency)*

T *(Term)*	**DF*(Document Frequency)***	**D/DF**	**IDF*(Inverse Document Frequency)***
Keren	1	3	log 3 = 0.477
Banget	2	1.5	log 1.5 = 0.176
Ikut	1	3	log 3 = 0.477
Magang	1	3	log 3 = 0.477

Setelah mengetahui nilai IDF *(Inverse Document Frequency)* langkah selanjutnya langsung dapat menghitung TF-IDF, adapun hasil dari contoh di atas dapat dilihat pada Tabel 6.

Tabel 6. Contoh perhitungan TF-IDF

Q	TF			DF	D/DF	IDF	IDF+1	W = TF*(IDF+1)		
	D1	D2	D3					D1	D2	D3
Keren	0	1	0	1	3	0.477	1.477	0	1477	0
Banget	1	1	0	2	1.5	0.176	1.176	1176	1176	0
Ikut	0	0	1	1	3	0.477	1.477	0	0	1477
Magang	0	0	1	1	3	0.477	1.477	0	0	1477
Nilai Bobot Setiap Dokumen								1176	2653	2954

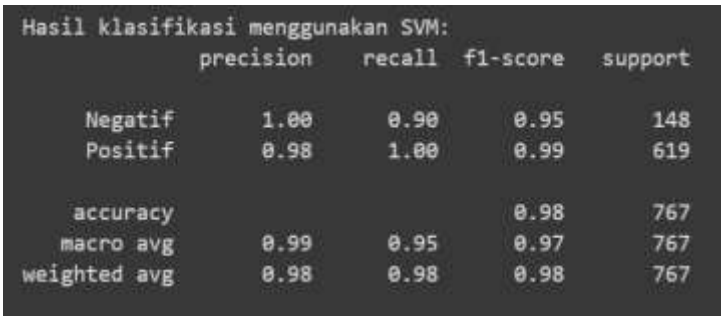
Adapun hasil pada word vector yang sudah terbobot dapat dilihat pada Tabel 7.

Tabel 7. Contoh dari *Word Vector* yang sudah terbobot

	Keren	Banget	Ikut	Magang
(D1)	0	1176	0	0
(D2)	1477	1176	0	0
(D3)	0	0	1477	1477

2. Impementasi Klasifikasi SVM

Untuk menerapkan analisis sentimen dalam penelitian ini, menggunakan metode Support Vector Machine (SVM) sebagai algoritma klasifikasi. Library Scikit-learn, yang merupakan salah satu library paling populer untuk pembelajaran mesin di Python, digunakan untuk mengimplementasikan SVM. Peneliti menggunakan fungsi train_test_split untuk memisahkan dataset menjadi subset pelatihan dan uji dengan proporsi 80:20. Variabel X_tfidf merupakan vektor fitur TF-IDF yang digunakan sebagai fitur masukan, sedangkan variabel y menyimpan label sentimen yang ditetapkan.



Gambar 3. Hasil klasifikasi SVM

Akurasi sebesar 0.98 menunjukkan sejauh mana model SVM mampu mengklasifikasikan sentimen secara keseluruhan dengan benar. Akurasi yang tinggi menunjukkan bahwa model SVM memberikan hasil yang akurat dalam memprediksi sentimen pada dataset yang digunakan.

3. Implementasi Klasifikasi Naïve Bayes

Selain metode SVM, penulis juga menerapkan metode Naive Bayes sebagai algoritma klasifikasi untuk analisis sentimen. Digunakan library Scikit-learn untuk mengimplementasikan Naive Bayes.

Pertama, membagi dataset menjadi subset pelatihan dan uji menggunakan fungsi train_test_split dengan proporsi 80:20. Variabel X_tfidf merupakan vektor fitur TF-IDF yang digunakan sebagai fitur masukan, dan variabel y menyimpan label sentimen."

Hasil klasifikasi menggunakan Naive Bayes:

	precision	recall	f1-score	support
Negatif	1.00	0.20	0.34	148
Positif	0.84	1.00	0.91	619
accuracy			0.85	767
macro avg	0.92	0.60	0.63	767
weighted avg	0.87	0.85	0.80	767

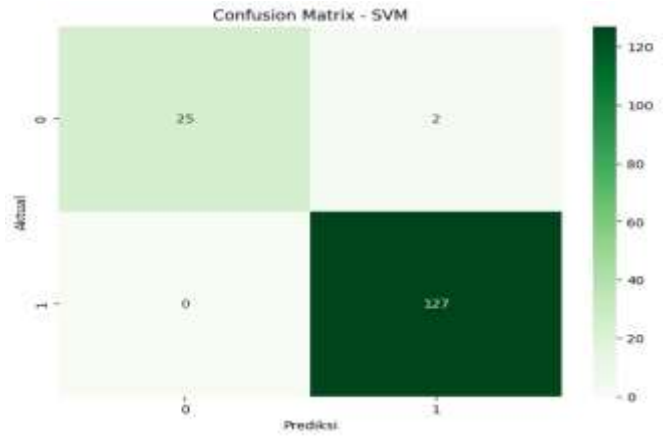
Gambar 4. Hasil Klasifikasi Naïve Bayes

Akurasi sebesar 0.85 menunjukkan sejauh mana model Naive Bayes mampu mengklasifikasikan sentimen secara keseluruhan dengan benar. Meskipun akurasi tersebut tidak tinggi, hal ini dapat disebabkan oleh tantangan dalam mengklasifikasikan sentimen "Negatif" yang memiliki nilai *recall* yang rendah.

4. Evaluasi Model

Digunakan *Confusion matrix* untuk mengevaluasi kinerja sebuah model klasifikasi. Ini adalah gambar tabel yang menggambarkan hasil prediksi model dibandingkan dengan nilai sebenarnya dari dataset yang digunakan. *Confusion matrix* terdiri dari empat metrik utama: *true positive (TP), false positive (FP), true negative (TN), dan false negative (FN)*. Berikut adalah Gambar tabel *Confusion Matrix* dari SVM dan Naïve Bayes beserta penjelasan.

Confusion Matrix untuk Model SVM :



Gambar 5. Confusion Matrix SVM

Berikut adalah penjelasan mengenai confusion matrix SVM pada gambar diatas :

a. *True Positive (TP):* Jumlah dokumen yang benar diprediksi sebagai positif. Dalam kasus ini, TP adalah 127.

b. *True Negative (TN):* Jumlah dokumen yang benar diprediksi sebagai negatif. Dalam kasus ini, TN adalah 25.

c. *False Positive (FP):* Jumlah dokumen yang salah diprediksi sebagai positif. Dalam kasus ini, FP adalah 2.

d. *False Negative (FN):* Jumlah dokumen yang salah diprediksi sebagai negatif. Dalam kasus ini, FN adalah 0.

Menggunakan nilai-nilai ini, penulis dapat menghitung metrik evaluasi sebagai berikut:

a. Akurasi *(Accuracy):* Rasio dokumen yang diprediksi dengan benar (TN + TP) dengan jumlah total dokumen.

$$Accuracy = \frac{25+127}{25+2+0+127} = 0.985$$

b. Presisi *(Precision)*: Rasio dokumen positif yang diprediksi dengan benar (TP) dengan total dokumen yang diprediksi sebagai positif (TP + FP).

$$Precision = \frac{127}{127+2} = 0.984$$

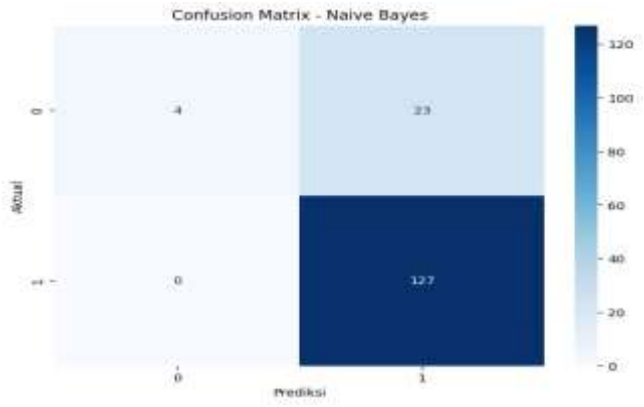
c. *Recall (Sensitivity):* Rasio dokumen positif yang diprediksi dengan benar (TP) dengan total dokumen positif sebenarnya (TP + FN).

$$Recall = \frac{127}{127+0} = 1.000$$

d. *F1-Score*: Ukuran gabungan dari presisi dan *recall*, menggunakan rumus:

$$F1\text{-}Score = 2 \times \frac{(0.984\ x\ 1.000)}{(0.984 + 1.000)} = 0.992$$

Confusion Matrix untuk Model Naive Bayes:



Gambar 6. Confusion Matrix Naïve Bayes

Berikut adalah penjelasan mengenai confusion matrix Naïve Bayes pada gambar diatas:

a. *True Positive (TP):* Jumlah dokumen yang benar diprediksi sebagai positif. Dalam kasus ini, TP adalah 127.
b. *True Negative (TN):* Jumlah dokumen yang benar diprediksi sebagai negatif. Dalam kasus ini, TN adalah 4.
c. *False Positive (FP):* Jumlah dokumen yang salah diprediksi sebagai positif. Dalam kasus ini, FP adalah 23.
d. *False Negative (FN):* Jumlah dokumen yang salah diprediksi sebagai negatif. Dalam kasus ini, FN adalah 0.

Menggunakan nilai-nilai ini, penulis dapat menghitung metrik evaluasi sebagai berikut:

a. Akurasi *(Accuracy):* Rasio dokumen yang diprediksi dengan benar (TN + TP) dengan jumlah total dokumen.

$$Accuracy = \frac{4+127}{4+23+0+127} = 0.857$$

b. Presisi *(Precision):* Rasio dokumen positif yang diprediksi dengan benar (TP) dengan total dokumen yang diprediksi sebagai positif (TP + FP).

$$Precision = \frac{127}{127+23} = 0.847$$

c. Recall *(Sensitivity)*: Rasio dokumen positif yang diprediksi dengan benar (TP) dengan total dokumen positif sebenarnya (TP + FN).

$$Recall = \frac{127}{127+0} = 1.000$$

d. *F1-Score*: Ukuran gabungan dari presisi dan recall, dihitung menggunakan rumus:

$$F1\text{-}Score = 2 \times \frac{(0.847\ x\ 1.000)}{(0.847 + 1.000)} = 0.917$$

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil klasifikasi menggunakan metode SVM dan Naive Bayes pada dataset yang digunakan, dapat disimpulkan sebagai berikut:

a. Berdasarkan analisis yang dilakukan pada data dari media sosial Twitter, sebagian besar pandangan masyarakat terhadap implementasi program Kampus Merdeka cenderung positif.
b. Akurasi model SVM 98%, menunjukkan kemampuan yang baik dalam memprediksi sentimen secara keseluruhan pada dataset yang digunakan.
c. Naive Bayes menunjukkan hasil yang menarik dalam mengklasifikasikan sentimen, dengan nilai presisi yang baik untuk kedua kelas "Negatif" dan "Positif", dan mendapatkan akurasi model sebesar 85 %.
d. F1-score untuk kelas "Negatif" dapat ditingkatkan, menunjukkan tantangan dalam mengklasifikasikan dokumen dengan sentimen "Negatif".

Saran

Dari hasil penelitian yang dilakukan masih banyak kekurangan, dengan demikian peneliti berharap penelitian ini untuk dikembangkan, beberapa saran dari penulis, di antaranya :

a. Penanganan ketidakseimbangan kelas: Jika terdapat ketidakseimbangan antara jumlah sampel dalam kelas sentimen, menerapkan teknik penanganan ketidakseimbangan seperti oversampling atau undersampling dapat membantu meningkatkan kinerja model dalam mengklasifikasikan sentimen yang kurang representatif.
b. Melakukan analisis yang lebih mendalam pada klasifikasi sentimen negatif dan mengidentifikasi faktor-faktor yang mempengaruhi performa model dapat membantu meningkatkan kinerja model dalam mengklasifikasikan sentimen negatif yang lebih kompleks atau ambigu.

DAFTAR PUSTAKA

[1] E. Suhartini, A. Yumarni, and S. Maryam, “Pelaksanaan Program Merdeka Belajar Kampus Merdeka dalam upaya peningkatan kinerja perguruan tinggi,” *DIDAKTIKA TAUHIDI: Jurnal Pendidikan Guru Sekolah Dasar*, vol. 9, no. 1, pp. 65–78, 2022.

[2] E. Simatupang and I. Yuhertiana, “Merdeka belajar kampus merdeka terhadap perubahan paradigma pembelajaran pada pendidikan tinggi: Sebuah tinjauan literatur,” *Jurnal Bisnis, Manajemen, Dan Ekonomi*, vol. 2, no. 2, pp. 30–38, 2021.

[3] R. A. Wati, H. Irsyad, and M. E. Al Rivan, “Klasifikasi Pneumonia Menggunakan Metode Support Vector Machine,” *J. Algoritm*, vol. 1, no. 1, pp. 21–32, 2020.

[4] H. Apriyani and K. Kurniati, “Perbandingan Metode Naïve Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus,” *Journal of Information Technology Ampera*, vol. 1, no. 3, pp. 133–143, 2020.

[5] H. Hermanto, A. Mustopa, and A. Y. Kuntoro, “Algoritma Klasifikasi Naive Bayes Dan Support Vector Machine Dalam Layanan Komplain Mahasiswa,” *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, vol. 5, no. 2, pp. 211–220, 2020.

[6] S. J. Pipin and H. Kurniawan, "Analisis Sentimen Kebijakan MBKM Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM," *Jurnal SIFO Mikroskil*, vol. 23, no. 2, pp. 197–208, 2022.

[7] R. Apriani and D. Gustian, "Analisis Sentimen dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia," *Jurnal Rekayasa Teknologi Nusa Putra*, vol. 6, no. 1, pp. 54–62, 2019.

[8] R. A. Raharjo, I. M. G. Sunarya, and D. G. H. Divayana, "Perbandingan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Kasus Analisis Sentimen Terhadap Data Vaksin Covid-19 Di Twitter," *Elkom: Jurnal Elektronika dan Komputer*, vol. 15, no. 2, pp. 456–464, 2022.

[9] I. C. R. Drajana, "Metode support vector machine dan forward selection prediksi pembayaran pembelian bahan baku kopra," *ILKOM Jurnal Ilmiah*, vol. 9, no. 2, pp. 116–123, 2017.

[10] L. Fregoso-Aparicio, J. Noguez, L. Montesinos, and J. A. García-García, "Machine learning and deep learning predictive models for type 2 diabetes: a systematic review," *Diabetol Metab Syndr*, vol. 13, no. 1, p. 148, 2021.

[11] E. Alfonsius and M. Rifai, "PERANCANGAN SISTEM INFORMASI PENJUALAN BARANG BERBASIS VENDOR MANAGED INVENTORY (VMI)," *PROSIDING SEMANTIK*, vol. 1, no. 2, p. 253, 2015.

[12] M. Rifai, E. Alfonsius, and L. Sanjaya, "PEMODELAN SISTEM INFORMASI ALUMNI STMIK ADHI GUNA BERBASIS WEBSITE," *SEMNASTEKNOMEDIA ONLINE*, vol. 5, no. 1, pp. 1–2, 2017.

[13] K. P. Simanjuntak and U. Khaira, "Pengelompokkan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering: Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 7–16, 2021.

[14] G. Landeras, A. Ortiz-Barredo, and J. J. López, "Forecasting weekly evapotranspiration with ARIMA and artificial neural network models," *Journal of irrigation and drainage engineering*, vol. 135, no. 3, pp. 323–334, 2009.

[15] A. A. Adebiyi, A. O. Adewumi, and C. K. Ayo, "Comparison of ARIMA and artificial neural networks models for stock price prediction," *J Appl Math*, vol. 2014, no. 1, p. 614342, 2014.

[16] R. Ramadhan, Y. A. Sari, and P. P. Adikara, "Perbandingan Pembobotan Term Frequency-Inverse Document Frequency dan Term Frequency-Relevance Frequency terhadap Fitur N-Gram pada Analisis Sentimen," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 11, pp. 5075–5079, 2021.